CLAIMS

What is claimed is:

1) In an Internet domain name server (DNS), a method of providing an Internet Protocol (IP) address of a first one of a plurality of servers of an Internet host, the method comprising the steps of:

receiving a DNS resolution request; and

returning the IP address of the first one of the plurality of servers such that a total number of times that the IP address of the first one of the plurality of servers is returned is proportional to a relative weight of the first one of the plurality of servers relative to a total weight of the plurality servers.

2) The method described in claim 1 including the additional steps of:

periodically determining whether each one of the plurality of servers is down;

setting to zero the relative weight of each one of the plurality of servers that is down; and

recalculating the total weight of the plurality of servers.

3) The method described in claim 2 including the additional step of informing the Internet host of each one of the plurality of servers that is down.

4) The method described in claim 1 including the additional steps of:

periodically determining whether each one of the plurality of servers is generating a time out error;

reducing the relative weight of each one of the plurality of servers that generate the time out error; and

recalculating the total weight of the plurality of servers.

5) The method described in claim 4 including the additional step of notifying the Internet host of each one of the plurality of servers that is generating the time out error.

6) The method described in claim 1 wherein the first one of the plurality of servers is returned in a list including IP addresses of the other ones of the plurality of servers such that a total number of times that the IP address of the first one of the plurality of servers is at a top of the list is proportional to the relative weight of the first one of the plurality of servers relative to the total weight of the plurality servers.

7) In a domain name server (DNS), a method of providing an Internet Protocol (IP) address of one of a plurality of servers of an Internet host, the method comprising the steps of:

identifying a plurality of client domains that have recently requested to be connected to the plurality of servers;

apportioning the plurality of client domains among the plurality of servers such that a percentage of requests served by each one of the plurality of servers is proportional to a relative weight of each respective one of the plurality of servers;

receiving a DNS resolution request from one of the plurality of client domains;

returning an IP address of one of the plurality of servers to which the one of the client domains was apportioned.

8) The method described in claim 7 wherein the step of apportioning the plurality of client domains among the plurality of servers includes the step of identifying a number of times each one of the plurality of client domains has recently requested being connected to the plurality of servers.

9) The method described in claim 8 wherein the step of identifying a number of times each one of the plurality of client domains has

recently requested being connected to the plurality of servers includes the step of reading a recent server log of the Internet host.

10) The method described in claim 7 wherein the IP address of the one of the plurality of servers to which the one of the client domains was apportioned is returned in a list including IP addresses of the other ones of the plurality of servers such that the IP address of the one of the plurality of servers to which the one of the client domains was apportioned is at a top of the list.

11) The method described in claim 7 including the additional steps of:

periodically determining whether each one of the plurality of servers is down;

setting to zero the relative weight of each one of the plurality of servers that is down; and

recalculating the total weight of the plurality of servers.

12) The method described in claim 11 including the additional step of notifying the Internet host of each one of the plurality of servers that is down.

13) The method described in claim 7 including the additional steps of:

periodically determining whether each one of the plurality of servers is generating a time out error;

reducing the relative weight of each one of the plurality of servers that generate the time out error; and

recalculating the total weight of the plurality of servers.

14) The method described in claim 13 including the additional step of notifying the Internet host of each one of the plurality of servers that is generating the time out error.

15) The method described in claim 7 including the additional steps of:

receiving a DNS resolution request from a client domain not included in the plurality of client domains that have recently requested to be connected to the plurality of servers; and

returning the IP address of a first one of the plurality of servers such that a total number of times that the IP address of the first one of the plurality of servers is returned is proportional to a leftover capacity of the first one of the plurality of servers relative to a total leftover capacity of the plurality servers.

16) The method described in claim 7 wherein the plurality of client domains that have recently requested to be connected to the plurality of servers include client domains that have recently frequently requested to be connected to the plurality of servers.

17) In a domain name server (DNS), a method of providing an Internet Protocol (IP) address of one of a plurality of servers of an Internet host, the method comprising the steps of:

determining a relative weight for each one of the plurality of servers;

calculating a total weight of the plurality of servers

identifying a plurality of client domains that have recently frequently requested to be connected to the plurality of servers;

identifying a number of times each one of the plurality of client domains has recently connected to the plurality of servers

calculating a total number of times the plurality of servers has been connected to the plurality of client domains;

apportioning the plurality of client domains among the plurality of servers such that the relative weight of the first one of the plurality of servers in comparison to the total weight of the plurality servers is approximately proportional to a cumulative number of times each one of the plurality of client domains apportioned to the first one of the plurality of servers has recently connected to the plurality of servers in comparison to the total

number of times the plurality of servers has been connected to the plurality of client domains;

receiving a DNS resolution request from one of the plurality of client domains; and

returning an IP address of one of the plurality of servers to which the one of the client domains was apportioned.

18) The method described in claim 17 including the additional steps of:

receiving a DNS resolution request from a client domain not included in the plurality of client domains that have recently frequently requested to be connected to the plurality of servers; and

returning the IP address of a first one of the plurality of servers such that a total number of times that the IP address of the first one of the plurality of servers is returned is proportional to a leftover capacity of the first one of the plurality of servers relative to a total leftover capacity of the plurality servers.

19) The method described in claim 17 including the additional steps of:

periodically determining whether each one of the plurality of servers is down;

setting to zero the relative weight of each one of the plurality of servers that is down; and

recalculating the total weight of the plurality of servers.

20) The method described in claim 17 including the additional steps of:

periodically determining whether each one of the plurality of servers is generating a time out error;

reducing the relative weight of each one of the plurality of servers that generate the time out error; and

recalculating the total weight of the plurality of servers.

Add A3